PREDICTING SELF-LEADERSHIP IN MILITARY CADETS: A MACHINE LEARNING APPROACH

António Maria Rosinha Dias Barbosa, Center for Leadership Development Studies, Military Academy, barbosa.amrd@exercito.pt

DOI: https://doi.org/10.60746/8_16_42458

ABSTRACT

Self-leadership is a key competency in military contexts, influencing decision-making, adaptability, and operational performance. This study proposes a data-driven approach to optimise the measurement of self-leadership among cadets at a military academy.

Building on the original theoretical framework by Neck and Manz (1992, 1996), and the measurement model developed by Houghton and Neck (2002), which comprises nine factors and 35 items, we applied Recursive Backward Elimination combined with Artificial Neural Networks in MATLAB to iteratively reduce the number of predictive variables.

After six elimination rounds, an optimised model with four factors and 18 items was obtained, achieving a predictive accuracy of $R^2 \approx 0.93$ with only a minimal increase in mean squared error. This streamlined version retained 92.6% of the explained variance while significantly reducing the evaluative burden.

Comparisons between the original and optimised models, based on classification accuracy and Pearson correlation, revealed high consistency, with the optimised version – hereafter referred to as Mod_4F_RBE – producing the fewest discrepancies. This confirms the robustness of the reduced structure for both continuous and categorical applications.

These results highlight the potential of machine learning to optimise psychometric instruments and promote their integration into applied social science. The proposed methodology offers a replicable model for improving assessment efficiency in demanding environments.

Keywords: Self-Leadership, Artificial Neural Networks, Recursive Backward Elimination, Machine Learning, Leadership Prediction.

1. INTRODUCTION

The growing complexity of military operations and leadership demands has renewed interest in intrapersonal competencies such as self-regulation, autonomy, and adaptability. Among these, self-leadership – introduced by Manz (1986) and later expanded by Neck and Manz (1992, 1996) – is defined as the process through which individuals influence themselves to achieve the self-direction and motivation required for effective performance. Within military education and training, it has gained prominence due to its role in developing autonomous leaders capable of operating under pressure.

The Revised Self-Leadership Questionnaire (RSLQ), developed by Houghton and Neck (2002), is a key instrument for measuring this construct, encompassing nine factors across three core dimensions. Its validity has been demonstrated across diverse cultural and professional contexts (Houghton, Dawley, & DiLiello, 2012; Carmeli et al., 2006). However, Ho and Nesbit (2014) questioned whether such comprehensive tools could be streamlined without compromising psychometric quality.

Simultaneously, the rise of machine learning (ML) in the social and behavioural sciences has enabled new approaches to refining theoretical models (Bzdok, Altman, & Krzywinski, 2018). Among these techniques, Artificial Neural Networks show particular promise in modelling complex, non-linear relationships often beyond the reach of traditional methods. Yet, their application in leadership research remains limited, partly due to concerns over accessibility and interpretability (Lepri et al., 2017).

Recursive Backward Elimination (RBE), a feature selection technique that iteratively removes less impactful variables while monitoring model performance, is particularly promising. Widely used in fields such as bioinformatics and finance (Guyon & Elisseeff, 2003), its application to psychometric optimisation is still in early stages. Preliminary evidence (e.g., Zhou et al., 2021) suggests its potential to identify minimal predictor sets while preserving explanatory power.

Building on these developments, this study explores the potential of ML – specifically Artificial Neural Networks and RBE – to optimise self-leadership measurement in a military context. Using a validated dataset of cadet responses to the RSLQ, it aims to: (1) identify the most salient predictors of self-leadership, and (2) assess whether a reduced-factor model derived through ML can maintain robust predictive performance. The research was guided by two questions:

- Which self-leadership factors best predict overall self-leadership among military cadets?
- Can ML techniques – namely Artificial Neural Networks and Recursive Backward Elimination – generate a more parsimonious self-leadership model without compromising psychometric robustness?

The remainder of this article is structured as follows. Section 2 reviews the theoretical foundations of self-leadership, the RSLQ structure, and ML applications in the social sciences. Section 3 describes the methodology, including dataset preparation and RBE implementation. Section 4 presents and discusses the results of the model reduction process. Section 5 outlines the conclusions, practical implications, and future directions.

2. LITERATURE REVIEW

2.1 SELF-LEADERSHIP

The concept of self-leadership was introduced by Charles Manz (1986) and subsequently developed by Christopher Neck and colleagues (Houghton & Neck, 2002; Neck & Houghton, 2006). It refers to individuals' capacity to influence their own behaviour, thoughts, and emotions in pursuit of personal and professional goals. The construct is structured around three dimensions: behaviour-focused strategies, natural reward strategies, and constructive thought pattern strategies. These are operationalised through nine factors, typically measured using the 35-item Revised Self-Leadership Questionnaire (RSLQ) (Houghton & Neck, 2002).

Self-leadership distinguishes itself from related constructs such as self-regulation and self-management by emphasising proactive and intentional self-influence. Whereas self-regulation typically involves reactive control in response to external stimuli, self-leadership represents a more internalised, future-oriented form of personal agency. Empirical studies have shown that it contributes uniquely to outcomes such as motivation, performance, and well-being, even when controlling for other self-regulatory processes (Neck & Houghton, 2006; Stewart et al., 2011).

The RSLQ has been widely validated across educational, healthcare, and organisational settings. Its factorial structure has been replicated in multiple countries, supporting its robustness and cross-cultural relevance (Houghton et al., 2012; Ho & Nesbit, 2014). This validation strengthens its position as a foundational tool for assessing self-leadership, especially in structured environments such as military education.

To simplify the RSLQ while preserving its theoretical core, Houghton et al. (2012) proposed the Abbreviated Self-Leadership Questionnaire (ASLQ), a nine-item version reflecting the three core dimensions. Though reliable and valid across professional settings, the ASLQ prioritises conceptual representation over empirical optimisation, leaving open the possibility that alternative approaches – particularly those focused on predictive performance – may yield more parsimonious and analytically robust models.

Table 1 summarises the nine RSLQ factors, grouped by their respective theoretical dimensions, along with a brief description of each factor's core function.

Table 1

Factors of the RSLQ grouped by theoretical dimensions

Dimension	Factor / Variable	Brief Description
Behavior-Focused Strategies	1. Goal Setting (SelfGoal)	Promotion of goal-setting and strategies for enhancing personal performance (Manz, 1986; Neck & Manz, 2010).
	2. Self-Reward (SefRew)	Self-feedback strategies that reinforce desired behaviours (Houghton & Neck, 2002).
	3. Self-Punishment (SelfPun)	Self-feedback strategies aimed at correcting ineffective behaviours (Houghton & Neck, 2002; Neck & Houghton, 2006).
	4. Self-Observation (SelfObs)	Monitoring one's own behaviour and performance. (Houghton & Neck, 2002; Neck & Houghton, 2006).
	5. Self-Cueing (SelfCue)	Use of visual/auditory cues to support performance (Houghton & Neck, 2002; Neck & Houghton, 2006).
Natural Reward Strategies	6. Natural Reward Strategies (NatRew)	Task reconfiguration to increase intrinsic satisfaction and reduce negative perceptions (Houghton & Neck, 2002; Neck & Houghton, 2006).
Constructive Thought Patterns	7. Visualizing Successful Performance (VisSuccPerf)	Mental rehearsal of success scenarios before undertaking a task (Houghton & Neck, 2002; Neck & Houghton, 2006).
	8. Self-Talk (SelfTalk)	Development of positive internal dialogue to support motivation (Houghton & Neck, 2002; Neck & Houghton, 2006).
	9. Evaluating Beliefs and Assumptions (EvaBelAss)	Critical analysis of dysfunctional beliefs that interfere with performance (Burns, 1980; Ellis, 1977).

Note. The abbreviations presented here (e.g., SelfGoal, NatRew, VisSuccPerf) are used throughout the manuscript to refer to the corresponding self-leadership factors.

2.2 MACHINE LEARNING IN THE SOCIAL SCIENCES

The following section addresses how ML techniques are being applied to optimise the assessment and modelling of psychological constructs.

Modern organisations operate in increasingly complex environments marked by limited resources and rapid technological change. In this context, Decision Support Systems (DSS) have evolved to assist managers by integrating data, models, and intuitive interfaces. The integration of Artificial Intelligence (AI) further enhances these systems, enabling reasoning under uncertainty using both empirical and theoretical knowledge.

Within AI, Case-Based Reasoning aids decision-making through analogies with past cases. More broadly, ML has gained prominence in the social sciences for modelling complex human phenomena such as behaviour, leadership, and performance. Algorithms like ANNs, Decision Trees, and Support Vector Machines offer strong predictive power and generalisability across diverse datasets (Bzdok et al., 2018; Yarkoni & Westfall, 2017).

Studies by Whelan et al. (2020) and Khalil et al. (2022) showed that ML can classify leadership styles and predict performance using psychometric data. While still underused in military contexts, ML shows increasing promise in enhancing soft skills assessment and uncovering latent psychological traits.

Effective application of ML in psychology requires a structured predictive modelling process. Two key components are feature selection: identifying relevant input variables and modelling non-linear relationships. RBE supports the former by iteratively simplifying the model while preserving accuracy. ANNs, meanwhile, are well suited to capture complex interactions that traditional models often overlook. Despite their success in other fields (Guyon & Elisseeff, 2003), such techniques remain underexplored in leadership and psychometric research, where they offer considerable untapped potential.

2.3 RECURSIVE BACKWARD ELIMINATION WITH ARTIFICIAL NEURAL NETWORKS

RBE, also known as Recursive Feature Elimination, is a widely used ML technique for optimising models by identifying subsets of input variables that maximise predictive performance (Guyon et al., 2002). The process begins with the full set of variables and iteratively removes the least impactful one—i.e., the variable whose exclusion results in the smallest decrease in performance, typically assessed via Mean Squared Error (MSE) or the coefficient of determination (R^2).

When combined with ANNs, RBE is particularly effective in modelling complex, multidimensional constructs, such as those encountered in psychology and behavioural sciences. Although still underutilised in the social sciences, its application is growing as a method for refining and simplifying psychometric instruments (e.g., Abdar et al., 2021).

In addition to identifying key predictors, RBE facilitates the development of more efficient and interpretable models without compromising predictive validity. As such, it complements traditional approaches like confirmatory factor analysis and structural equation modelling, contributing to the advancement of quantitative leadership research.

3. MATERIALS AND METHODS

3.1 METHODS

The study employed an iterative approach to factor elimination, guided by the performance of ANN models developed in MATLAB's Neural Network Toolbox. Similar to reverse backward elimination, the aim was to identify the smallest set of predictive variables capable of retaining the explanatory power of the original model. In each iteration, one latent factor was removed and changes in performance were recorded, focusing on MSE and R^2, as recommended by Haykin (2009) and adopted in psychometric research (Zhang et al., 2019; Saghafi et al., 2021).

This methodology aligns with structural refinement strategies in ML, particularly relevant for psychometric optimisation without compromising predictive validity (Huang et al., 2022). The rationale is that factors with limited predictive contribution can be excluded, enhancing parsimony and practical applicability in contexts such as military training.

By comparing each model to its immediately preceding version, the approach enables controlled performance assessment and identifies the threshold beyond which model robustness declines – reflecting the logic proposed by Guyon and Elisseeff (2003) in their foundational work on feature selection.

RBE was used to reduce input factors while maintaining predictive accuracy. A baseline model with nine factors was trained in MATLAB's Neural Network Fitting Tool, using the Levenberg–Marquardt algorithm and a single hidden layer of 10 neurons.

At each iteration, one factor was excluded and the model retrained. The factor whose removal caused the smallest change in performance – measured by MSE and R^2 – was eliminated from subsequent iterations. The process continued until further elimination increased MSE by more than 5% or reduced R^2 by more than 2%, following thresholds suggested in feature selection literature (Guyon et al., 2002; Abdar et al., 2021).

Model performance was evaluated across training, validation, and testing sets using a 70/15/15 data split. For each iteration, MSE, R^2, and the number of training epochs were recorded to ensure robust comparisons between reduced models.

3.2 DATA COLLECTION AND PREPARATION

Data were collected at the Portuguese Military Academy in two separate waves. The first took place in the 2015/2016 academic year and yielded 311 valid responses; the second occurred in 2020/2021, with 366 valid responses. In both instances, data were gathered from the entire population of undergraduate cadets (1st to 4th year), ensuring full institutional coverage. Participation was voluntary, and responses were handled with strict confidentiality, allowing for linkage with academic performance indicators solely for research purposes and in line with institutional data handling protocols.

The dataset comprises responses from 677 military cadets who completed the 35-item RSLQ, developed by Houghton and Neck (2002). This instrument assesses nine self-leadership factors, grouped into three core dimensions: behaviour-focused strategies, natural reward strategies, and constructive thought pattern strategies.

Items were rated on a Likert scale and aggregated by factor. Factor scores were computed as the mean of their items and weighted proportionally to the number of items in each factor (e.g., a factor with five items received a weight of 5/35 = 0.1429). A global self-leadership index was then calculated by summing the weighted factor scores and normalising the result, yielding a final score between 0.36 and 1.8. This composite score served as the output variable in all ML models.

3.3 TOOLS AND SOFTWARE

All analyses were conducted using MATLAB R2015a® with the Neural Network Toolbox. Data preprocessing and preliminary calculations were performed in Microsoft Excel®. Pearson correlation analyses comparing the global self-leadership indices from the original and reduced (four-factor) models were conducted in IBM SPSS®.

4. PRESENTATION AND DISCUSSION OF RESULTS

4.1 SELF-LEADERSHIP MODEL VIA RECURSIVE BACKWARD ELIMINATION

The study began with an analysis of the complete self-leadership model proposed by Houghton and Neck (2002), which consists of nine factors and 35 items. Using the RBE algorithm, the objective was to identify the factors with the least predictive power in estimating the global self-leadership index. This evaluation was based on two primary criteria: variations in MSE and changes in R^2, measured after individually removing each factor. The global self-leadership index was calculated as a weighted mean of the factors, with weights reflecting the number of items associated with each factor.

The initial model, which included all nine factors, demonstrated near-perfect predictive performance ($R = 0.99999$; $MSE = 7.20391e^{-11}$) and served solely as the reference for the first iteration. Thereafter, each iteration used the reduced model from the previous round as its new base, ensuring progressive control over performance degradation throughout the optimization process. In each round, the factor whose removal yielded the smallest increase in MSE was deemed the least impactful and therefore excluded, following best practices in feature selection (Guyon & Elisseeff, 2003).

In the first iteration, the exclusion of SelfGoal produced the lowest increase in MSE ($8.70596e^{-5}$), leading to its removal.

In the second iteration, starting from an eight-factor model, SelfRew was excluded after showing the smallest impact on MSE ($2.14744e^{-4}$).

During the third iteration, the model still retained strong predictive power, and the removal of SelfPun was selected (MSE = $5.80576e^{-4}$).

The fourth iteration began with six factors, and SelfCue was excluded based on minimal impact (MSE = $1.31038e^{-3}$).

In the fifth iteration, SelfTalk was identified for exclusion (MSE = $1.26511e^{-3}$) from the remaining five-factor model.

Finally, the sixth iteration worked with a four-factor model, where any further removal caused unacceptable degradation of predictive performance ($R^2 = 0.926$, baseline MSE = $1.31038e^{-3}$).

Thus, the RBE process confirmed that the optimal model was reached after six iterations, preserving statistical robustness with only 18 items.

Final Predictive Model: The RBE process identified the following four factors as essential for maintaining predictive integrity:

- SelfPun (Self-punishment strategies);
- NatRew (Natural rewards);
- VisSuccPerf (Visualizing successful performance);
- EvaBelAss (Evaluating beliefs and assumptions).

Table 2 summarizes the progression of MSE values and the eliminated factors in each iteration.

Table 2

Mean squared error progression and factor removal across RBE iterations

	08 Var	07 Var	06 Var	05 Var	04 Var	03 Var
Fator removed	Training Mean Squared Error (MSE)					
SelfGoal	8,70596e-5	2,75507e-5				
SelfRew	1,62485e-4	2,14744e-4	3,74882e-4			
SelfPun	2,17516e-4	3,28731e-4	5,80576e-4	7,33337e-4	1,10319e-3	1,80379e-3
SelfObs	5,69430e-5					
SelfCue	1,13050e-4	1,44552e-4	4,58158e-4	6,90554e-4		
NatRew	2,09677e-3	2,80874e-3	3,24725e-3	3,59725e-3	4,31695e-3	5,01618e-3
VisSuccPerf	2,88446e-4	4,24198e-4	7,96433e-4	1,25766e-3	1,45136e-3	2,46096e-3
SelfTalk	4,13805e-4	4,08387e-4	7,15411e-4	7,58314e-4	1,31038e-3	
EvaBelAss	4,33042e-4	5,20485e-4	6,65568e-4	1,47952e-3	1,28941e-3	2,28462e-3

Note. Mean Squared Error (MSE) of the full model (with all variables) = 7,20391e-11

The RBE process, guided by the minimization of the MSE at each step, enabled the reduction of the construct from nine to four factors while preserving 92.6% of the variance explained by the original model. The retaining factors proved sufficient to robustly represent the self-leadership index, now composed of 18 items. The criterion of selecting the factor associated with the lowest MSE in each round demonstrated to be an objective and statistically consistent methodology, aligned with best practices in feature selection within ML (Guyon & Elisseeff, 2003; Jain & Zongker, 1997).

4.2 COMPARISON OF SELF-LEADERSHIP INDEXES IN THE FOUR-FACTOR MODEL

Based on the optimized four-factor model, a comparative analysis was conducted to assess how different weighting strategies impact the computation of the self-leadership index. Two versions of the index were calculated: one using the original theoretical weights derived from the number of items per factor (Mod_4F_Ori), and the other using weights based on the relative contribution of each factor as observed in the sixth and final round of RBE, namely the ΔMSE values (Mod_4F_RBE). This analysis aimed to determine which approach better preserves the structure and interpretability of the original construct while aligning with empirical performance.

The development of a reduced self-leadership index from the optimized model follows principles consistent with feature importance in ML. Specifically, the empirical contribution of each retained factor — expressed through its ΔMSE during the RBE process — was normalized and used as a relative weight. This weighting method reflects the actual impact of each factor on model performance and is widely applied in advanced model optimization practices (Molnar, 2022; Kuhn & Johnson, 2019). Similar strategies are found in feature engineering, model pruning, and ensemble methods such as Random Forests, Gradient Boosting, and neural networks (Breiman, 2001; Guyon & Elisseeff, 2003).

By anchoring the index in empirical importance, the Mod_4F_RBE version incorporates not only the most predictive dimensions but also the relative strength of their influence. This results in a more parsimonious yet informative model, which retains interpretability and predictive utility. Furthermore, the use of ΔMSE as a basis for weighting ensures that the resulting index remains psychometrically meaningful, integrating both theoretical and data-driven insights into self-leadership measurement.

4.2.1 QUALITATIVE ANALYSIS

To assess the consistency between the original nine-factor self-leadership index and the indexes generated by the optimized four-factor model, a categorical comparison was conducted. Each participant's self-leadership score was classified into three ordinal levels: low, medium, and high.

Using the original theoretical weighting structure, the reduced four-factor model produced 27 mismatches across 677 cases, corresponding to an error rate of 3.99%. In comparison, the optimised model weighted according to the empirical impact of each factor – derived from the ΔMSE values in the final RBE iteration and hereafter referred to as Mod_4F_RBE – produced only 21 classification discrepancies, yielding an error rate of 3.10%.

These results indicate a high degree of agreement between the full and reduced models. Moreover, Mod_4F_RBE exhibited slightly greater consistency with the original construct in terms of categorical classification. This suggests that, for practical applications involving classification or triage (e.g., training diagnostics or developmental feedback), the empirically optimised model offers both predictive precision and categorical reliability.

4.2.2 QUANTITATIVE ANALYSIS

To complement the categorical assessment, a Pearson correlation analysis was performed to quantify the alignment between the different versions of the self-leadership index. Specifically, correlations were calculated among:

(1) The original nine-factor model (Mod_Ori);

(2) The four-factor model using theoretical weights (Mod_4F_Ori);

(3) The four-factor model with empirical weights based on ΔMSE values (Mod_4F_RBE).

Table 3

Pearson correlation table

		Mod Orig	Mod 4F Orig	Mod 4F RBE
Mod Orig	Pearson Correlation	1	,877**	,934**
	Sig. (2-tailed)		,000	,000
	N	677	677	677
Mod_4F_Ori	Pearson Correlation	,877**	1	,785**
	Sig. (2-tailed)	,000		,000
	N	677	677	677
Mod_4F_RBE	Pearson Correlation	.934**	.785**	1
	Sig. (2-tailed)	,000	,000	
	N	677	677	677

Note. **. Correlation is significant at the 0.01 level (2-tailed).

The results revealed very strong correlations between the original model and both reduced versions: $r = .877$ for the model using theoretical weights (Mod_4F_Ori) and $r = .934$ for the empirically optimised model (Mod_4F_RBE) ($p < .001$ in both cases). These findings suggest that the reduced structure effectively preserves the core of the self-leadership construct. Notably, Mod_4F_RBE exhibited the highest correlation with the original index, indicating superior performance in terms of predictive precision – further validating the utility of ML techniques in identifying the most impactful predictors.

Although both reduced models retain substantial explanatory power, Mod_4F_RBE shows slightly greater numerical fidelity to the original construct. It also produced fewer classification discrepancies, reinforcing the idea that empirical weighting enhances both predictive accuracy and categorical reliability. Consequently, Mod_4F_RBE emerges as the most robust and versatile solution for practical applications, combining parsimony, precision, and theoretical coherence.

4.2.3 COMPARATIVE CONCLUSION OF THE FOUR-FACTOR SELF-LEADERSHIP INDEXES

The data indicate that for classification purposes – such as diagnosis, screening, or identifying training needs – the empirically optimised model (Mod_4F_RBE) is more suitable, as it yielded fewer discrepancies in categorizing individuals into "low," "medium," or "high" self-leadership levels. In contrast, the model using the original theoretical weights (Mod_4F_Ori), while conceptually aligned with the initial structure, resulted in a higher classification error.

When the objective is to obtain a continuous and more precise estimate of the self-leadership index – for instance, in individual monitoring or longitudinal tracking – both models demonstrate strong correlations with the full structure. However, Mod_4F_RBE offers superior predictive robustness and numerical fidelity.

Accordingly, Mod_4F_RBE emerges as the most effective and versatile solution for applied contexts. These findings illustrate how traditional psychometric frameworks can be successfully combined with contemporary ML techniques, enabling the constructive optimisation of assessment tools while preserving theoretical coherence.

4.3 SYNTHESIS OF THE INTERPRETATION OF RESULTS

As outlined in Table 2, the factors SelfPun, NatRew, VisSuccPerf, and EvaBelAss emerged as the strongest predictors of self-leadership. Systematically eliminating the remaining five factors reduced model complexity from 35 to 18 items, while preserving 92.6% of the variance explained by the original construct—demonstrating that explanatory power was largely retained despite scale reduction.

Comparative analyses of self-leadership index calculations showed that the empirically optimised model (Mod_4F_RBE) produced the fewest classification discrepancies and the highest correlation with the full structure. This indicates that Mod_4F_RBE offers superior performance both in categorical classification and in continuous predictive estimation.

These findings confirm the utility of Recursive Backward Elimination (RBE) in the social sciences. The method enabled the construction of a shorter, efficient, and empirically validated instrument, with applicability in educational and organisational settings. It proved not only computationally sound but also capable of optimising psychometric tools without compromising theoretical coherence.

Compared to the nine-item ASLQ developed by Houghton et al. (2012), which condenses the RSLQ along its original dimensions, the present four-factor model (Mod_4F_RBE) retains more predictive power (R^2 = 0.926) by leveraging a performance-driven ML approach. This highlights the added value of empirical feature selection in preserving the explanatory richness of the original construct.

In sum, this study introduces a refined self-leadership model that remains faithful to its theoretical foundations while addressing practical challenges in psychological assessment. The proposed Mod_4F_RBE structure is both psychometrically robust and operationally efficient, offering a viable alternative for agile and targeted evaluation.

5. CONCLUSION

This study investigated the application of ML techniques – specifically Recursive Backward Elimination (RBE) supported by Artificial Neural Networks (ANNs) – to optimise the assessment of self-leadership in a military training context. Starting from the original nine-factor structure grounded in the theoretical framework of Neck and Manz, the model was progressively reduced through six elimination rounds, ultimately yielding a streamlined version composed of four key factors and 18 items. Despite this substantial reduction, the optimised model preserved high predictive performance (R^2 = 0.926), confirming that it is possible to simplify the evaluative process without compromising measurement validity.

Comparative analyses between the original and reduced models further validated the new construct. When using empirical weights derived from the RBE procedure, the reduced model – hereafter referred to as Mod_4F_RBE – resulted in only 3.10% classification discrepancies and demonstrated a strong Pearson correlation with the full model ($r = .934$). These results underscore the robustness and fidelity of Mod_4F_RBE, supporting its use in contexts that require operational efficiency without sacrificing psychometric rigour.

Beyond the immediate findings, this research illustrates the growing relevance of ML methods in the social and behavioural sciences. By integrating computational modelling with classical psychometric frameworks, this study offers a replicable and empirically grounded pathway for construct optimisation. Such integration enhances both theoretical robustness and practical utility, particularly in domains such as leadership assessment and training design.

Nonetheless, some limitations must be acknowledged. The sample, although substantial ($n = 677$), is drawn from a single military academy, which may constrain the generalisability of the findings. Additionally, while the optimisation process prioritised predictive accuracy and parsimony, it did not yet address issues such as longitudinal stability or cross-cultural invariance. Future studies are encouraged to test the applicability of Mod_4F_RBE across diverse contexts, populations, and over time.

In conclusion, this work proposes a practical and scientifically robust tool – Mod_4F_RBE – for assessing self-leadership in military environments. It also contributes to a broader methodological dialogue, demonstrating how ML techniques can be meaningfully employed to enhance the development, refinement, and applicability of psychological instruments in real-world settings.

REFERENCES

Abdar, M., Pourpanah, F., Hussain, S., Rezazadegan, D., Liu, L., Ghavamzadeh, M., Fieguth, P., Cao, X., Khosravi, A., Acharya, U. R., Makarenkov, V., & Nahavandi, S. (2021). A review of uncertainty quantification in deep learning: Techniques, applications and challenges. *Information Fusion, 76*, 243–297. https://doi.org/10.1016/j.inffus.2021.05.008

Burns, D. D. (1980). *Feeling good: The new mood therapy*. William Morrow.

Breiman, L. (2001). *Random forests*. Machine Learning, 45(1), 5–32. https://doi.org/10.1023/A:1010933404324

Bzdok, D., Altman, N., & Krzywinski, M. (2018). Statistics versus machine learning. *Nature Methods, 15*(4), 233–234. https://doi.org/10.1038/nmeth.4642

Carmeli, A., Meitar, R., & Weisberg, J. (2006). Self-leadership skills and innovative behavior at work. *International Journal of Manpower, 27*(1), 75–90. https://doi.org/10.1108/01437720610652853

Ellis, A. (1977). *Rational emotive behavior therapy*. Springer Publishing Company.

Fernández-Delgado, M., Cernadas, E., Barro, S., & Amorim, D. (2014). Do we need hundreds of classifiers to solve real world classification problems? *Journal of Machine Learning Research*, *15*(1), 3133–3181. http://www.jmlr.org/papers/volume15/delgado14a/delgado14a.pdf

Guyon, I., & Elisseeff, A. (2003). An introduction to variable and feature selection. *Journal of Machine Learning Research*, *3*, 1157–1182. http://www.jmlr.org/papers/volume3/guyon03a/guyon03a.pdf

Guyon, I., Weston, J., Barnhill, S., & Vapnik, V. (2002). Gene selection for cancer classification using support vector machines. *Machine Learning, 46*(1), 389–422. https://doi.org/10.1023/A:1012487302797

Haykin, S. (2009). *Neural Networks and Learning Machines* (3rd ed.). Pearson.

Huang, H., Chen, X., & Wu, D. (2022). Feature selection and model optimization in social science research using neural networks. *Computational Intelligence and Neuroscience, 2022*, Article ID 9439285. https://doi.org/10.1155/2022/9439285

Houghton, J. & Neck, C. (2002). The Revised Self-Leadership Questionnaire: Testing a hierarchical factor structure for self-leadership. *Journal of Managerial Psychology*, *17*(8) 672-691. https://doi.org/10.1108/02683940210450484

Houghton, J. D., Dawley, D., & DiLiello, T. C. (2012). The abbreviated self-leadership questionnaire (ASLQ): A more concise measure of self-leadership. *International Journal of Leadership Studies, 7*(2), 216–232.

Lepri, B., Staiano, J., Sangokoya, D., Letouzé, E., & Oliver, N. (2017). *The tyranny of data? The bright and dark sides of data-driven decision-making for social good*. In *Transparent Data Mining for Big and Small Data* (pp. 61–75). Springer. https://doi.org/10.1007/978-3-319-54024-5_5

Khalil, R., Iqbal, R., Matulevičius, R., & Valli, C. (2022). A machine learning approach towards the prediction of leadership style using personality traits. *Expert Systems with Applications, 193*, 116491. https://doi.org/10.1016/j.eswa.2021.116491

Kuhn, M., & Johnson, K. (2019). *Feature engineering and selection: A practical approach for predictive models*. CRC Press.

Manz, C. C. (1986). Self-leadership: Toward an expanded theory of self-influence processes in organizations. *Academy of Management Review, 11*(3), 585–600. https://doi.org/10.5465/amr.1986.4306232

Molnar, C. (2022). *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable* (2nd ed.). https://christophm.github.io/interpretable-ml-book/

Neck, C. P., & Houghton, J. D. (2006). Two decades of self-leadership theory and research: Past developments, present trends, and future possibilities. *Journal of*

Managerial Psychology, 21(4), 270–295. https://doi.org/10.1108/02683940610663097

Neck, C. P., & Manz, C. C. (1992). Thought self-leadership: The influence of self-talk and mental imagery on performance. *Journal of Organizational Behavior, 13*(7), 681–699. https://doi.org/10.1002/job.4030130705

Neck, C. P., & Manz, C. C. (2010). *Mastering self-leadership: Empowering yourself for personal excellence* (5th ed.). Pearson Prentice Hall.

Pudil, P., Novovičová, J., & Kittler, J. (1994). Floating search methods in feature selection. *Pattern Recognition Letters*, *15*(11), 1119–1125. https://doi.org/10.1016/0167-8655(94)90127-9

Saghafi, F., Abedini, M., & Hosseini, M. (2021). Artificial neural network for modeling behavioral science constructs: A case study on resilience and decision-making styles. *Heliyon*, *7*(5), e06932. https://doi.org/10.1016/j.heliyon.2021.e06932

Whelan, E., Teigland, R., Donnellan, B., & Golden, W. (2020). How artificial intelligence and machine learning will transform project management. *International Journal of Project Management, 38*(3), 167–177. https://doi.org/10.1016/j.ijproman.2020.01.007

Yarkoni, T., & Westfall, J. (2017). Choosing prediction over explanation in psychology: Lessons from machine learning. *Perspectives on Psychological Science, 12*(6), 1100–1122. https://doi.org/10.1177/1745691617693393

Zhang, C., Zhao, T., & Geng, Y. (2019). A novel variable selection method for psychological measurement using artificial neural networks. *Frontiers in Psychology, 10*, 1458. https://doi.org/10.3389/fpsyg.2019.01458

Zhang, Y., & Zhang, S. (2019). Machine learning models in social science research: Toward interpretability, replicability, and reproducibility. *Political Analysis, 27*(2), 180–201. https://doi.org/10.1017/pan.2019.2

Zhou, Y., Lee, D., Jin, X., & Liu, X. (2021). A review on feature selection for data mining. *Computational Intelligence and Neuroscience, 2021*, 1–19. https://doi.org/10.1155/2021/6674663